NOAA Technical Memorandum NWS SR-91

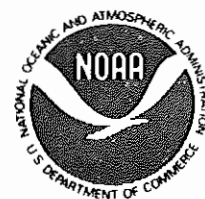SPECIAL PROGRAM FORECAST VERIFICATION - TWO CONTRIBUTIONS

AN EVALUATION OF CERTAIN AGRICULTURAL FORECAST PARAMETERS

Kenneth A. Wigner, WSFO, Lubbock, Texas

COMPUTERIZED VERIFICATION OF FIRE WEATHER FORECASTS

Donald L. Ocker, WSFO, Albuquerque, New Mexico

Scientific Services Division
Southern Region
Fort Worth, Texas
August 1977

AN EVALUATION OF CERTAIN AGRICULTURAL FORECAST PARAMETERS

Kenneth A. Wigner, WSFO, Lubbock, Texas

AN EVALUATION OF CERTAIN AGRICULTURAL

FORECAST PARAMETERS

by

Kenneth A. Wigner
National Weather Service Forecast Office
Lubbock, Texas

INTRODUCTION

Forecasters at WSFO, Lubbock routinely prepare agricultural forecasts twice
daily, in the early morning and late afternoon. These include predictions of
maximum and minimum relative humidities (RH) and amount (percentage) of possible
sunshine (SS). All RH forecasts are for 12-hour periods: 12GMT (00GMT) to
00GMT (12GMT) for a morning (afternoon) issuance time. In the morning forecast
package RH is forecast for three periods while in the afternoon this extends
through four periods, beginning with the maximum expected (tonight) and ending
with the minimum for the "day after tomorrow." The morning forecast for SS
includes "today" and tomorrow" while in the afternoon the predictions are for
"tomorrow" and the "day after tomorrow."

Although verification has been performed on some other forecast elements –
e.g., temperature and precipitation – no comprehensive, on-going scheme has
existed for RH and SS. It is the purpose of this paper to document forecast
performance on these parameters for the calendar year 1975.
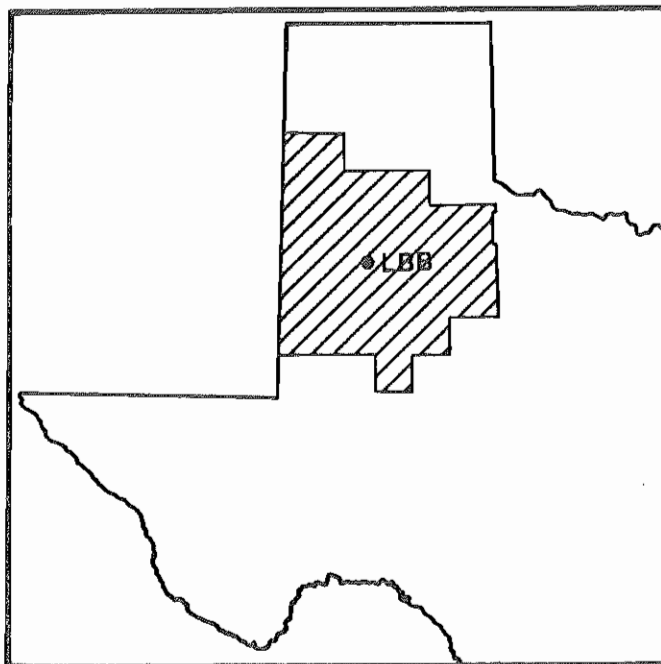
The agricultural forecasts are con-
sidered valid for a South Plains
area comprising 28 counties and
covering about 22,000 square miles
(see Figure I). The actual values
used for verification were those
observed at Lubbock. Admittedly,
verifying area-type forecasts
with point observations introduces
possible ambiguities; but these
were the ones available, and con-
sidering that Lubbock is centrally
located and a substantial data
sample was at hand the likely degree
of representativeness was deemed
acceptable.

Figure I. Area (hatched) for which
agricultural forecasts considered
valid.

VERIFICATION METHODS AND RESULTS

The forecasts were evaluated by two methods. The first was devised on-station in an effort to obtain useful results; it is used for individual forecasters as well as for the office product. The second method employs a more conventional approach, with contingency tables yielding <u>bias</u>, <u>percent correct</u>, etc.; it is used only for overall staff performance.

Tables 1-6 exhibit the results from the first method. All forecasts during 1975 for both RH and SS were in terms of deciles - i.e., 10-20%, 20-30%, etc. For the purposes of this method of verification each forecast was assigned the central value of the decile - e.g., 60-70% = 65%. A positive (negative) error will be taken to mean that the forecast value was greater (less) than the verifying observation. For instance, a SS forecast of, say, 40-50% (= 45%) of possible sunshine versus an observed 73% would yield a <u>negative</u> error of 73-45 = 28.

As seen in the tables, positive and negative errors were accumulated separately and jointly for both elements and average absolute errors were computed for all stratifications of the forecasts. The average algebraic error is also included to provide a measure of the tendency to over- or underforecast. An interesting sidelight not apparent in the tables was the relatively pronounced tendency of forecasters with long experience in the South Plains area to forecast too much sunshine and too little moisture. Note in Tables 1 and 4 that individuals are identified by number and ranked by average absolute error for RH and SS separately. Therefore, forecaster No. 1 for RH is not necessarily No. 1 in the SS ranking.

Results of the contingency table method of verification are summarized in Tables 7-10 and an example contingency table is shown in Appendix A. Appendix B explains the computation procedure for the skill scores, etc.

Relative Humidity

Table 2 shows staff performance on a monthly basis. The average absolute errors range from about 8 to 15 percent, with the overall tendency being distinctly toward forecasting RH too low (negative algebraic error). The forecasts were also analyzed by period and a summary for the year is shown in Table 3. Not unexpectedly, a systematic increase in average absolute error with increasing lead time was revealed - this is also reflected in Table 7 where decreasing skill scores are found. Table 7 also indicates that skill is very high at forecasting within one category of the verifying decile; i.e., large misses are infrequent. There does not appear to be a significant difference in staff ability to forecast a maximum or minimum RH.

Bias-by-category and average bias for each forecast period is shown in Table 8. Recall that a bias of 1.00 means that the category was forecast the same number of times that it was observed (not necessarily concurrently) and that values greater (less) than 1.00 indicate positive (negative) bias; i.e., the category was forecast more often (less often) than observed. Also, average absolute error by period on a monthly basis is depicted in Figures 2a and 2b.

<u>Sunshine</u>

SS is somewhat more difficult to forecast than is RH. Relatively, however, average absolute errors are generally lower during the warm season (as with RH), due largely to more persistent conditions. Average station errors for the year ranged from about 11 to 25 percent (see Table 5). On an annual basis, as seen in Tables 5 and 10, there is a distinct overall positive bias in SS forecasts - i.e., a tendency to forecast too much sunshine. Tables 6 and 9 reflect in two ways the expected decrease of skill with lead time. Note that average errors (skill scores) are higher (lower) for the second period than for the first for both morning and evening forecasts. Recall also that Period I for the evening forecast has a longer lead time than does Period I for the morning forecast.

Though not included here, the contingency tables from which the data in Tables 9 and 10 were derived made it evident that forecasters were quite reluctant to predict extremes of SS. There were few forecasts of less than 10% or greater than 90%.

Figures 3a and 3b depict average absolute error by period on a monthly basis.

SUMMARY AND CONCLUSIONS

A scheme for verifying forecasts of SS and RH has been devised and utilized on data for the calendar year 1975. The lower average errors occurred, as might be expected, during the warm season when weather conditions were more persistent. The more difficult of the two parameters to forecast successfully proved to be SS, which is essentially a forecast of cloudiness. A bias toward forecasting too dry and too sunny was identified.

Identification of individual biases should prove helpful to forecasters in improving these products. An expanded data sample may show other tendencies, hence this project will continue. A possible conclusion from the magnitude of the verification statistics is that more objective methods need to be developed to provide guidance for forecasts of SS and RH. Means for such development are presently being examined.

ACKNOWLEDGEMENTS

Forecaster Number	Number of Forecasts	Cumulative Positive Error	Cumulative Negative Error	Total Absolute Error	Average Absolute Error	Average Algebraic Error
1	113	424	546	970	8.6	-1.1
2	336	1864	1318	3182	9.5	1.6
3	306	1480	1461	2941	9.6	0.1
4	303	1154	1834	2988	9.9	-2.2
5	318	976	2165	3141	9.9	-3.7
6	112	575	601	1176	10.5	-0.2
7	145	650	902	1552	10.7	-1.7
8	355	1904	2082	3986	11.2	-0.5
9	141	602	1008	1610	11.4	-2.9
10	121	557	854	1411	11.7	-2.5
11	133	767	881	1557	11.7	-1.5
12	119	455	1056	1511	12.7	-5.5
Staff	2502	11317	14708	26025	10.4	-1.4

Table 1.   INDIVIDUAL FORECASTER VERIFICATION
FOR RELATIVE HUMIDITY (1975)

Month	Number of Forecasts	Cumulative Positive Error	Cumulative Negative Error	Total Absolute Error	Average Absolute Error	Average Algebraic Error
January	217	1153	1530	2683	12.7	-1.7
February	196	701	1480	2181	11.3	-4.0
March	217	1357	1230	2587	11.9	0.6
April	210	1516	1560	3076	14.7	-0.2
May	214	1054	934	1988	9.3	0.6
June	210	706	1138	1844	8.8	-2.1
July	201	514	1526	2040	10.2	-5.0
August	217	744	894	1638	7.6	-0.7
September	208	1059	656	1715	8.3	1.9
October	217	703	1061	1764	8.1	-1.7
November	210	1197	996	2193	10.4	1.0
December	185	613	1703	2316	12.5	-5.9
Year	2502	11317	14708	26025	10.4	-1.4

Table 2.   RELATIVE HUMIDITY VERIFICATION (1975)

Period	Number of Forecasts	Cumulative Positive Error	Cumulative Negative Error	Total Absolute Error	Average Absolute Error	Average Algebraic Error
MORNING						
1	360	1657	1690	3347	9.3	-0.1
2	360	1102	2345	3447	9.6.	-3.5
3	358	1959	2113	4072	11.4	-0.4
EVENING						
1	356	1183	1917	3100	8.7	-2.1
2	356	1726	1963	3689	10.4	-0.7
3	356	1629	2307	3936	11.1	-1.9
4	356	2061	2373	4434	12.5	-0.9
Year	2502	11317	14708	26025	10.4	-1.4

Table 3. RELATIVE HUMIDITY VERIFICATION BY PERIOD

Forecaster	Number of Forecasts	Cumulative Positive Error	Cumulative Negative Error	Total Absolute Error	Average Absolute Error	Average Algebraic Error
1	74	768	373	1141	15.4	5.3
2	168	2145	917	3062	18.2	7.3
3	72	843	501	1344	18.7	4.8
4	86	826	854	1680	19.5	-0.3
5	178	2514	969	3483	19.6	8.7
6	172	2038	1445	3483	20.3	3.5
7	186	1497	2326	3823	20.6	-4.5
8	190	1696	2234	3930	20.7	-2.8
9	68	1004	403	1407	20.7	8.8
10	90	967	1010	1977	22.0	-0.5
11	76	1067	613	1680	22.1	6.0
12	80	1304	558	1862	23.3	9.3
Staff	1440	16669	12203	28872	20.1	3.1

Table 4. INDIVIDUAL FORECASTER VERIFICATION
FOR SUNSHINE (1975)

Month	Number of Forecasts	Cumulative Positive Error	Cumulative Negative Error	Total Absolute Error	Average Absolute Error	Average Algebraic Error
January	124	1787	1215	3002	24.2	4.6
February	112	1356	1486	2842	25.4	-1.2
March	124	1428	1705	3133	25.3	-2.2
April	120	1507	1286	2793	23.3	1.8
May	124	1258	930	2188	17.7	2.7
June	120	1382	695	2077	17.3	5.7
July	122	2231	720	2951	24.2	12.4
August	124	1210	881	2091	16.9	2.7
September	120	835	1133	1968	16.4	-2.5
October	124	778	614	1392	11.2	1.3
November	118	1226	1002	2228	18.9	1.9
December	108	1671	536	2207	20.4	10.5
Year	1440	16669	12203	28872	20.1	3.1

Table 5.   SUNSHINE VERIFICATION (1975)

Period	Number of Forecasts	Cumulative Positive Error	Cumulative Negative Error	Total Absolute Error	Average Absolute Error	Average Algebraic Error
MORNING						
1	364	3353	2471	5824	16.0	2.4
2	364	4346	3149	7495	20.6	3.3
EVENING						
1	356	4093	2987	7080	19.9	3.1
2	356	4877	3596	8473	23.8	3.6
Year	1440	16669	12203	28872	20.1	3.1

Table 6.   SUNSHINE VERIFICATION BY PERIOD

Period	Number	R	C	S	R_1	C_1	S_1
MORNING							
1	360	143	40	.24	294	82	.77
2	360	130	36	.18	284	79	.73
3	358	113	32	.13	264	74	.67
EVENING							
1	356	143	40	.23	297	83	.78
2	356	124	35	.18	279	78	.73
3	356	114	32	.12	266	75	.67
4	356	103	29	.10	256	72	.65
ANNUAL	2502	870	35	.17	1976	79	.73

Table 7. RELATIVE HUMIDITY VERIFICATION (R indicates the number of forecasts falling within the correct decile, C is the percent of correct forecasts and S is the skill score. The addition of the subscript 1 indicates the inclusion of forecasts being within one category of the correct decile.)

Period	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	Average
MORNING											
1	.21	.96	1.31	1.07	1.31	.84	.77	.25	1.00	*	.78
2	*	*	*	.78	1.13	1.42	1.69	1.44	.78	.41	.77
3	.12	.93	1.49	1.01	1.24	.68	.62	.45	.67	*	.72
EVENING											
1	*	*	.50	.67	.44	1.74	1.31	1.19	.98	.56	.74
2	.08	1.19	1.25	.96	1.27	.83	.92	.45	*	*	.70
3	*	*	.50	.63	.40	1.20	1.56	1.36	.93	.49	.66
4	.12	1.03	1.30	1.14	1.22	.67	.77	.18	.33	*	.60

Table 8. RELATIVE HUMIDITY: BIAS-BY-CATEGORY AND AVERAGE BIAS (*indicates not forecast, not observed, or neither)

Period	Number	R	C	S	R_1	C_1	S_1
MORNING	364	140	38	.23	223	61	.51
1	364	113	31	.15	197	54	.43
2							
EVENING							
1	356	129	36	.20	209	59	.48
2	356	77	22	.05	163	46	.34
ANNUAL	1440	459	32	.16	792	55	.44

Table 9. SUNSHINE VERIFICATION (R indicates the number of forecasts falling within the correct decile, C is the percent of correct forecasts and S is the skill score. The addition of the subscript 1 indicates the inclusion of forecasts being within one category of the correct decile.)

Period	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	Average
MORNING											
1	.37	.38	2.00	.80	1.54	1.14	1.11	1.74	1.15	.87	1.11
2	.24	.23	.57	1.47	1.15	.86	1.69	1.78	1.46	.76	1.02
EVENING											
1	.36	*	1.29	.88	1.69	.95	1.64	1.26	1.59	.81	1.16
2	.15	.25	.86	1.14	1.07	1.00	1.77	1.85	1.86	.63	1.06

Table 10. SUNSHINE: BIAS-BY-CATEGORY AND AVERAGE BIAS (*indicates not forecast, not observed, or neither)
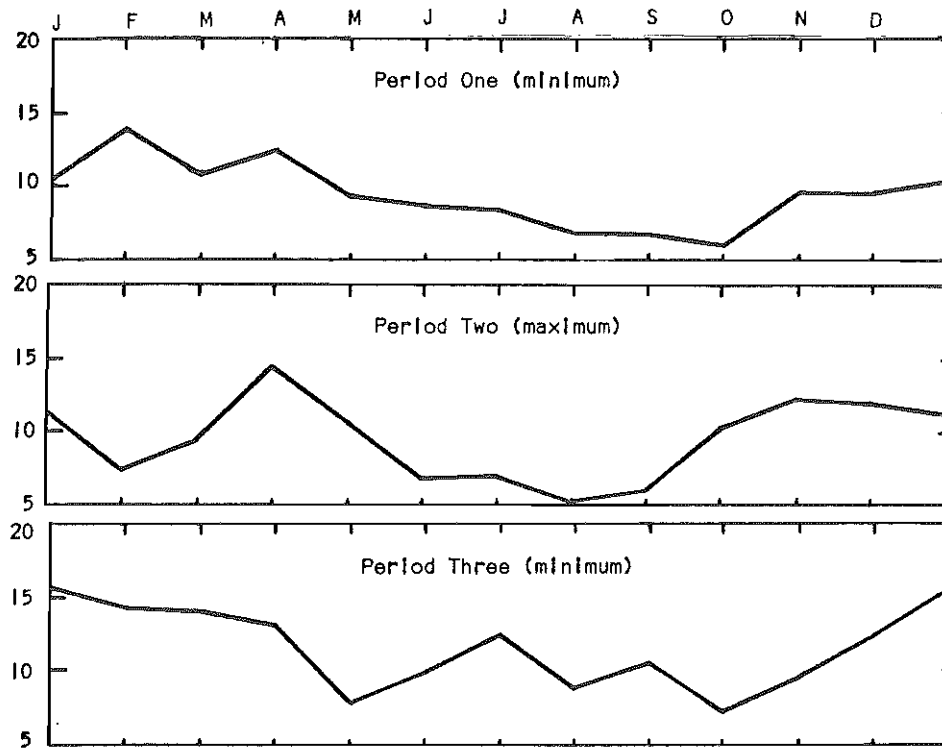
Fig. 2a. Average absolute
  error by month for
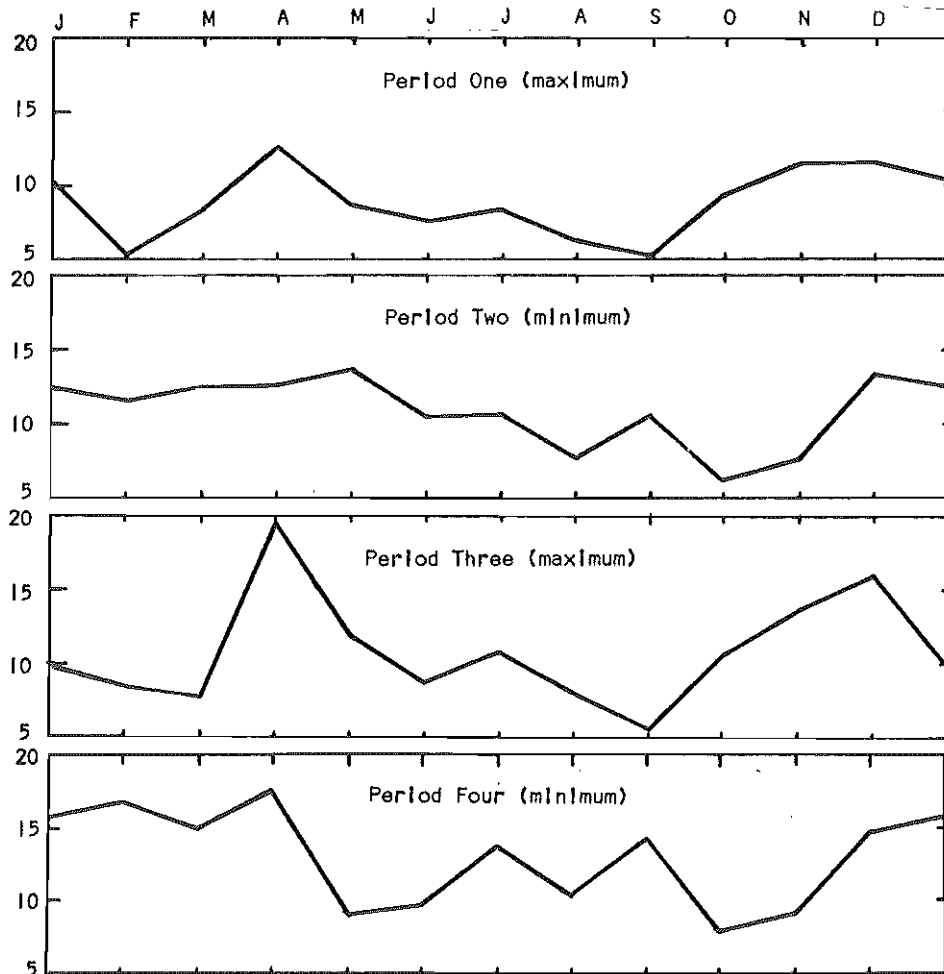 morning forecasts of
  relative humidity.

Fig. 2b. Average absolute
  error by month for
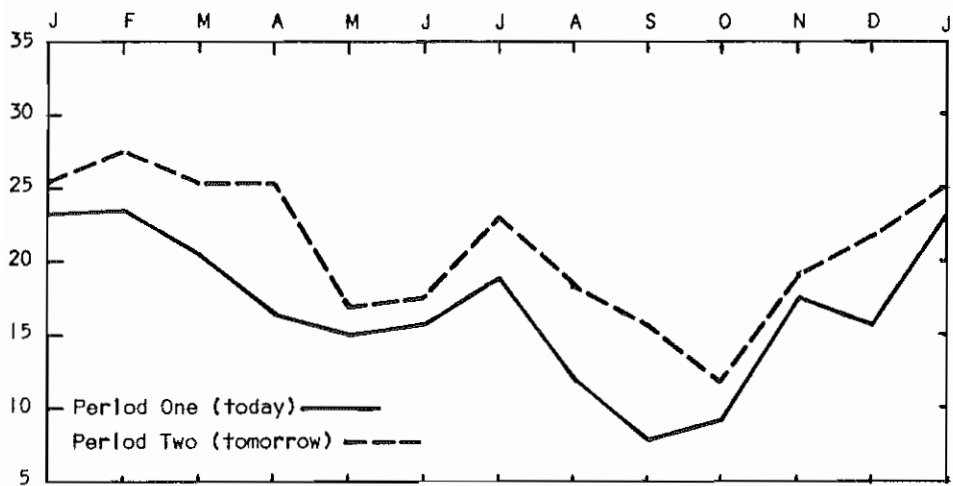 evening forecasts of
  relative humidity.

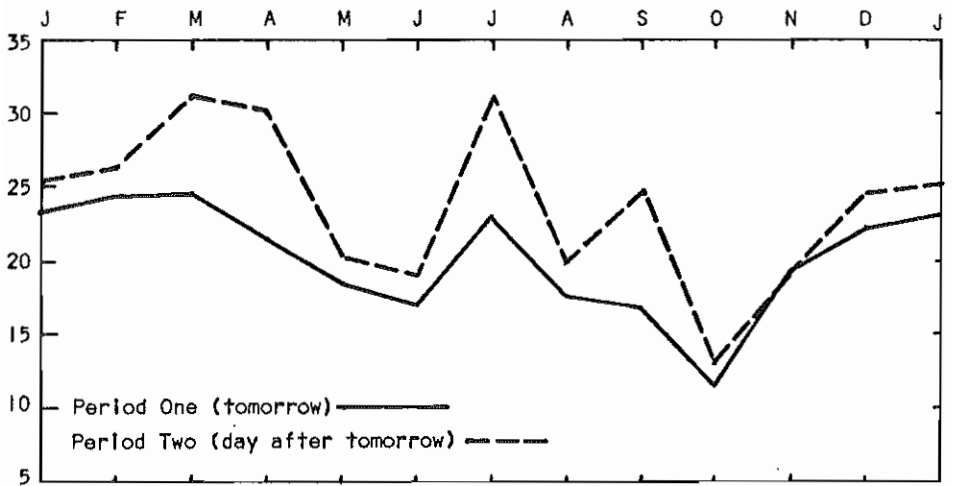Fig. 3a. Average absolute error by month for
morning forecasts of sunshine.

Fig. 3b. Average absolute error by month for
evening forecasts of sunshine.

O B S E R V E D

FORECAST		0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100	Total
	0-10	4		1								5
	11-20	15	42	14	4							75
	21-30	4	27	49	32	5	2	4	1			124
	31-40		7	24	37	12	6	3	1			90
	41-50	1	2	5	10	5	3	3	5			34
	51-60			1	1	4	3	2	1	1	3	16
	61-70			1			4		2		3	10
	71-80						1	1	1			3
	81-90								1	2		3
	91-100											
	Total	24	78	95	84	26	19	13	12	3	6	360 Grand Total

Example contingency table --
RH: 1st period, morning forecast

Appendix B

Development of Skill Score.

After Panofsky and Brier the expected number of correct forecasts based on chance (E_R) is computed from the following formula:

$$E_R = \frac{\Sigma R_i C_i}{T}$$

where R_i is the total of the i-th row and C_i is the total for the i-th column and T is the grand total. In my example:

5	X	24	=	120
78	X	75	=	5850
95	X	124	=	11780
84	X	90	=	7560
26	X	34	=	884
19	X	16	=	304
13	X	10	=	130
12	X	3	=	36
3	X	3	=	9
				26673

$E_R = \dfrac{26673}{360} = 74.1$

R = Number of correct forecasts

= 143

S = Skill Score

$= \dfrac{R - E_R}{T - E_R} = \dfrac{143 - 74.1}{360 - 74.1}$

S = .24

When we include one category either side of correct

$R_1 = 294 \qquad S_1 = \dfrac{294 - 74.1}{360 - 74.1} = .77$

Also, percent correct forecasts:

$C = \dfrac{R}{T} = \dfrac{143}{360} = 40\% \qquad C_1 = \dfrac{R_1}{T} = \dfrac{294}{360} = 82\%$

$\text{Bias} = \dfrac{\text{Number Forecast}}{\text{Number Observed}}$ (for each category)

NOAA Technical Memorandum NWS SR-91

COMPUTERIZED VERIFICATION OF FIRE WEATHER FORECASTS

Donald L. Ocker, WSFO, Albuquerque, New Mexico

Scientific Services Division
Southern Region
Fort Worth, Texas
August 1977

COMPUTERIZED VERIFICATION OF

FIRE WEATHER FORECASTS

by

Donald L. Ocker
National Weather Service Forecast Office
Albuquerque, New Mexico

INTRODUCTION

When the Interactive Real-Time Information System (IRIS)* became available
at WSFO, Albuquerque, one of the first tasks selected for computer appli-
cation was fire weather forecast verification. Computer capability at
Albuquerque seemed to present a way to alleviate the tedious data com-
pilation and repetitive calculations which are involved in verification
and are a deterrent to comprehensive verification. However, computers
usually accomplish only half the job since data compilation and entry
must still be performed manually in most cases. Use of computer-generated
eight-level paper teletype tapes of fire weather observations and fore-
casts has solved the hand labor problem. Verification has been accom-
plished with most data reduction, compilation and input done by machine.
Also, the data file utilized by the verification program can be used for
other jobs.

The Fire Weather System

Weather observations are taken daily by fire control personnel for the
calculation of fire danger using the National Fire Danger Rating System.
The system produces various indices relating to fire occurrence and behavior,
and which are required for fire control. In addition, forecast indices
are calculated using 24-hour weather forecasts supplied by the National
Weather Service (NWS).

These forecast indices are used for planning and staffing for the following
day.

In recent years, the fire danger calculations have been made by computer
using a U. S. Forest Service program called Administrative and Forest Fire
Retrieval and Management System (AFFIRMS). AFFIRMS is available to Forest
Service authorized users of the General Electric time-share computer system.

Daily observations from various locations are entered into AFFIRMS by fire
control personnel, who subsequently obtain a display of observed fire danger
indices. At most NWS offices, the forecaster enters AFFIRMS to obtain a
display of all observations in his area of interest. Twenty-four hour
weather forecasts are then prepared by the forecaster and entered into
AFFIRMS. This makes 24-hour forecasts of fire danger indices available
to fire control personnel when they again enter AFFIRMS and obtain a display.

*IRIS is a mini-computer system maintained by the Office of Technical Ser-
vices at National Weather Service Headquarters. The Southern Region uses
this system for local and regional studies through time-share terminals.
Programming is done in the BASIC language.

System Modifications in the Southwest

Fire control agencies in Arizona and New Mexico use a slightly different system. A teletype circuit using Model 33 ASR Teletypewriters links all control agencies with the WSFOs in Albuquerque and Phoenix. ASR-33s are used because they employ 8-level ASCII code which is compatible with most computer systems. Weather observations from locations in Arizona and New Mexico are transmitted via teletype to the Forest Service Regional Office in Albuquerque. There the observations are reperforated and entered into AFFIRMS. A display of observed fire danger indices is then requested from AFFIRMS, reperforated and transmitted to field offices via teletype. NWS forecasters use the original teletype observations to prepare forecasts which are transmitted via teletype to the Forest Service Regional Office for entry into AFFIRMS. A display of forest fire danger indices is then requested, reperforated and transmitted to field offices.

Although complete AFFIRMS weather displays are not available using the teletype system, displays of the fire danger indices include observed and forecast values of temperature, relative humidity and wind speed -- the primary variables required for fire weather forecast verification. Therefore, eight-level paper tapes containing all needed verification data are available via teletype or from the Forest Service Regional Office. All that is needed to use the data is an IRIS program to input the AFFIRMS tape, extract needed data and store the data in a verification file.

The Data Reduction Program

An IRIS program has been prepared which inputs the AFFIRMS tape line by line while writing into a work file. The display heading is located and the date and data-type (forecast or observation) is determined. The line of data for the first station is then located and values of temperature, relative humidity, wind speed, 10-hour timelag fuel moisture (10HR) and energy release component are extracted. Although not a weather variable, 10HR is included for verification since it is weather-related and thereby forecast directly or indirectly. Energy release component is dependent upon several variables; the most changeable is weather. Therefore, it is included as a possible check of overall forecast quality.

The program substitutes a recognizable constant if data for a particular station are not located. Also, routines are included to detect and correct garbled headings and to recognize forecasts under a heading for one month but valid for the first day of the following month. After needed values are extracted for each of the 22 fire danger stations in New Mexico, the data are stored in a verification file.

The verification file consists of 32 records, one for each day of the month plus an extra for the observed data of the last day of the previous month (needed to calculate observed 24-hour variability). Each record contains forecast values valid for the day, observed values for the day and a forecaster number which is added through use of another program.

The final result is a file consisting of both forecast and observed values of five variables for each day of the month. That is a total of up to 6851 individual values, all for the minimum effort required to run 62 paper tapes through the IRIS terminal.

File Uses

A verification program uses the file to verify each of the five variables. Verification results are produced for each individual fire danger station, for each of eight fire weather zones which contain groups of stations, and for all stations combined. Results are also produced for each individual forecaster. Results are by month and for the season to date. The verification consists of absolute error, bias, 24-hour variability and percent improvement over 24-hour variability (improvement over no forecast).

Although the data file was created as a verification data base, it is also valuable as a source of dependent samples to be used by a screening regression program.

RESULTS

Specific results are not presented here due to the very limited amount of data presently available. But previous AFFIRMS forecasts of 10HR have been unsatisfactory. This program uses a routine to calculate forecast 10HR from several forecast weather parameters, while observed 10HR is determined by weighing a set of half-inch ponderosa pine sticks (observed weight is compared with oven-dry weight to obtain percentage of moisture). Calculated forecasts of 10HR have not compared favorably with observed 10HR, either because of the AFFIRMS method or because of poor forecasts of the weather variables. This has led to development of a method for forecasting 10HR directly.

A screening regression program stored in IRIS was employed to relate observed 10HR to various observed and forecast meteorological parameters. Results thus far have been inconclusive since only mid-summer data have been examined. Better correlations may be found during the late spring and early summer fire season when large-scale changes should outweigh local effects.